THE LOGIC OF COMPARATIVE THEORY EVALUATION

by

M.L.G. Redhead

Chelsea College, London

ABSTRACT

Schaffner's logic of comparative theory evaluation is criticised for an inappropriate analysis of ad hocness. An alternative analysis, based on Zahar's account of novelty, is given and extended to the case of multiple successful predictions by a theory. The application of the method to the appraisal of quantitative prediction is discussed.

The Logic of Comparative Theory Evaluation *

1. The Bayesian Analysis of Ad hocness.

In a recent note Schaffner ([1974]) has given a formal discussion of the notion of ad hocness in terms of a Bayesian model for the appraisal of theories. Schaffner develops his general ideas in the context of a critique of Zahar's [1973] which was concerned with the particular problem of comparing the Einstein and Lorentz research programmes. Zahar suggests the following analysis of ad hocness[1]:

> Ad hoc_1: A theory is said to be ad hoc_1 if it has no novel consequences as compared with its predecessor.
>
> Ad hoc_2: [A theory]... is ad hoc_2 if none of its novel predictions have been actually 'verified'.
>
> Ad hoc_3: [A]... theory is said to be ad hoc_3 if it is obtained from its predecessor through a modification of the auxiliary hypotheses which does not accord with the spirit of the heuristic of the programme.

Zahar explains the meaning of novelty as follows[2]

> A fact will be considered novel with respect to a given hypothesis if it did not belong to the problem-situation which governed the construction of the hypothesis.

Schaffner begins his elucidation by discussing the notions of ad hoc_1 and ad hoc_3. The first he describes as a logical dream, since the novel consequences of a theory cannot in practice be "surveyed", so the question of whether a theory is ad hoc_1 can only be discussed relative to the extent to which novel consequences have been looked for at the particular epoch of the evaluation. Ad hoc_3 Schaffner claims to be "vague to the point of inapplicability". For Schaffner ad hoc_2 is "close to

---

* I am grateful to Donald Gillies, Jon Dorling and Noretta Koertge for helpful comments and suggestions. The present work formed part of a paper read in Professor Watkins' seminar at The London School of Economics in January 1976.

the sense in which ad hoc is used in science" and he announces that his Bayesian analysis will be brought to bear on this sense of ad hoc.

Denoting by $p(T/b \& e)$ the probability of a theory T to be true in the light of background knowledge b and the positive outcome e of some experiment not part of b, by $p(T/b)$ the prior assessment of T, by $p(e/T \& b)$ the probability of obtaining e given T and b, and by $p(e/b)$ the probability of obtaining e on the basis of background knowledge alone, Schaffner writes

$$p(T/b \& e) = \frac{p(T/b).p(e/T \& b)}{p(e/b)} \qquad (1)$$

If T explains e we can set $p(e/T \& b) = 1$ so in this case we obtain

$$p(T/b \& e) = p(T/b) \big/ p(e/b) \qquad (2)$$

Schaffner proceeds to discuss ad hocness as a property of an hypothesis as a constituent part of a theory, but in order to keep the argument as simple as possible we shall follow Zahar in considering the ad hocness of theories. Translated into these terms Schaffner's idea is that a theory T gives an ad hoc explanation of an experimental result e if $p(T/b)$ is close to zero and $p(e/b)$ is significantly larger than $p(T/b)$. The argument for $p(T/b)$ being small is that it has no "theoretical support" (or indeed empirical support other than e itself). This looks suspiciously like a reference to ad hoc_3, and Schaffner now appears to be claiming that a theory is ad hoc_2 partly in virtue of its being ad hoc_3. So he is not really giving an independent analysis of ad hoc_2 at all.

The lack of novelty in the prediction of e is associated by Schaffner with a high value for $p(e/b)$. On Zahar's account this is a necessary condition for lack of novelty, but not a sufficient condition. We proceed to show how using Zahar's notion

of novelty one can get an "internal" Bayesian analysis of ad hoc_2.

We express Bayes's theorem in the following familiar way

$$p(T/b \& e) = \frac{p(T/b).p(e/T\& b)}{p(e/T\& b)p(T/b) + p(e/\sim T\& b).p(\sim T/b)} \quad (3)$$

where $\sim T$ denotes the negation of T.

We follow Schaffner in interpreting p(A/B) as the degree of belief that A is true given that B is true. Note also that e is supposed in equation (3) to refer to a prediction made by the theory T (we could perhaps write e_T to emphasize this important point) so $p(e/\sim T\& b)$ means the probability that the prediction e_T derived from the theory T is a true prediction given that the background information b is true but the theory T is false (i.e., its consequences are not guaranteed to be true although "by accident" they may be true).

Writing $p(T/b) = x$, $p(e/\sim T\& b) = \varepsilon$

and taking $p(e/T\& b) = 1$ and using $p(\sim T/b) = 1 - x$

$$p(T/b\& e) = \frac{x}{x + \varepsilon(1 - x)} \quad (4)$$

We define an enhancement ratio γ by

$$\gamma = \frac{p(T/b\& e)}{p(T/b)}$$

whence using (4) we obtain the simple result

$$\gamma = \frac{1}{x + \varepsilon(1 - x)} \quad (5)$$

We can now explain that if a theory T is ad hoc_2 with respect to the experiment e then $\varepsilon = 1$, i.e. the explanation of e by T in no way depends on the truth or falsehood of T, both of which eventualities lead with certainty to the result e. This is just what a scientist means when he says T was an ad hoc explanation of e, namely T was devised for the express purpose of explaining e,

so the explanation of e is guaranteed independent of whether T is true or false. To show the consistency of our analysis if we put $\varepsilon = 1$ in (5) we get $\Upsilon = 1$, so the posterior and prior probabilities of T are equal (there is no enhancement) and this again is just what we expect from an ad hoc explanation of e, namely e itself gives us no additional information for assessing the truth of T.[3]

Notice that $p(e/b) = x + \varepsilon(1 - x)$ is equal to unity if $\varepsilon = 1$, but that $p(e/b) = 1$ is achieved for $x = 1$ whatever the value of ε, so the explication of novelty in terms of low $p(e/b)$ (Schaffner) and small ε (Zahar) are by no means equivalent.

On our analysis if $x \ll \varepsilon \ll 1$, then $\Upsilon \simeq 1/\varepsilon$, so in this case we get a big enhancement and the theory is far from being ad hoc. Noting that under these conditions $p(e/b) \simeq \varepsilon$, we see that this is a situation in which Schaffner would claim that the theory was ad hoc, which highlights the way in which his analysis differs from ours. Effectively Schaffner requires Υx to be small as his condition of ad hocness. He is thus concerned with the absolute value of the probability of a theory after it has explained some experimental finding. If this absolute value is still small the theory is to be regarded as an ad hoc explanation of the experiment. On our account the important aspect in assessing ad hocness for this case is not the absolute value of the probability but the enhancement ratio. Our point against Schaffner is not that his analysis may not explicate some legitimate sense of the ambiguous appellation ad hoc, but that it fails to explicate Zahar's very important notion of ad hoc$_2$. There is no inconsistency here on Schaffner's part, since he finds Zahar's account of ad hoc$_2$ inadequate in respect of the definition of novelty involved with its [redacted] historical associations, but we would maintain that Zahar's sense of ad hoc$_2$ is the one that ought to be explicated since it is the one most importantly used in science.

2. The Case of Multiple Predictions

To develop our analysis a little further we can consider how a theory builds up a favourable appraisal as it makes a number of successful predictions $e_1, e_2 \ldots e_n$ say. Denote by p_s the posterior probability $p(T/b \,\&\, e_1 \,\&\, e_2 \ldots \&\, e_s)$ after s successful predictions. Assuming the predictions are quite independent and for simplicity are all associated with the same value of ε, we can clearly write

$$p_n = \gamma^{(n)} \times \gamma^{(n-1)} \times \ldots \times \gamma^{(1)} \cdot p_o$$

where $\quad p_o = x$ according to our previous notation

and $\quad \gamma^{(s)} = \dfrac{1}{p_{s-1}(1-\varepsilon)+\varepsilon}$

$$p_{s+1} = \gamma^{(s+1)} \cdot p_s$$

The solution of this recursion is by inspection or more simply by replacing ε by ε^n in the formula for p_1 (see (4) above). We obtain[4]

$$p_n = \frac{1}{1-\varepsilon^n + \varepsilon^n/x} \qquad (6)$$

We can also ask what is the probability for the $(n+1)^{th}$ prediction e_{n+1} being correct if the theory has already made n successful predictions $e_1 \ldots e_n$. Denoting this probability by $p(e_{n+1})$ we clearly obtain the result

$$p(e_{n+1}) = \varepsilon + \frac{1-\varepsilon}{1-\varepsilon^n + \varepsilon^n/x} \qquad (7)$$

If ε is a small quantity (i.e. $\ll 1$) which will be the case if T is non-ad hoc_2 with respect to all the predictions, we can write the following formulae which will be perfectly satisfactory for the subsequent discussion

$$p_n \simeq \frac{1}{1+\varepsilon^n/x}, \qquad (8)$$

$$p(e_{n+1}) \simeq \mathcal{E} + p_n \qquad (9)$$

$$\gamma^{(n)} \simeq \frac{1}{\mathcal{E} + p_{n-1}} \qquad (10)$$

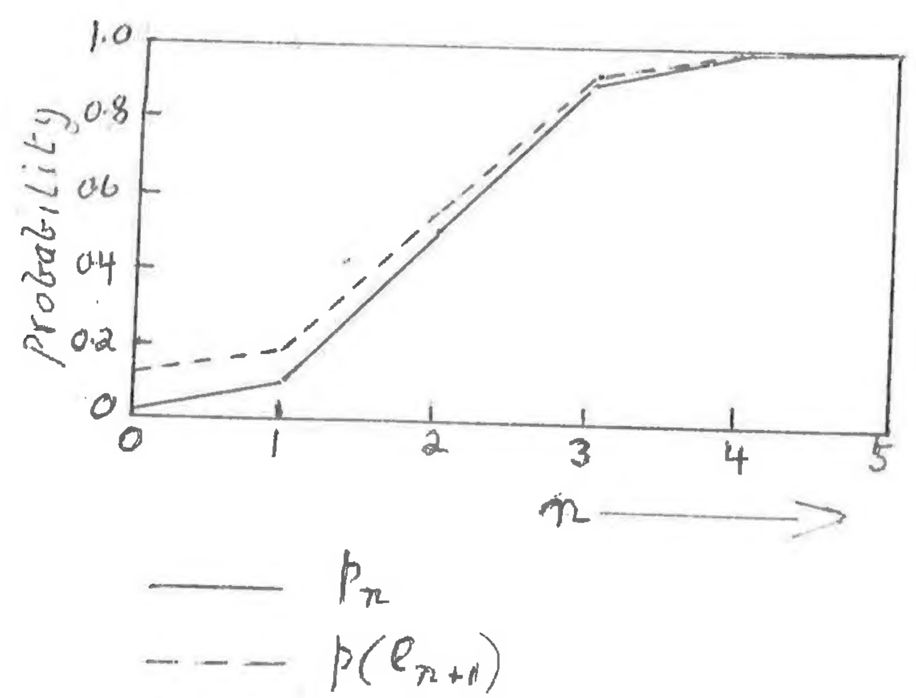
We note the following features

(1) The value of pn depends entirely on the ratio $\mathcal{E}^n/x$
For $\mathcal{E}^n/x \gg 1$, $p_n \ll 1$
and for $\mathcal{E}^n/x \ll 1$, $p_n \simeq 1$
At the critical point $\mathcal{E}^n = x$, we have $p_n \simeq 1/2$.
So if initially $x \ll \mathcal{E}$ as n increases pn will rise steeply as we reach the value $n = \ln x/\ln \mathcal{E}$.

(2) So long as $p_n \ll \mathcal{E}$, we have $p(\ell_{n+1}) \simeq \mathcal{E}$, but as pn builds up towards unity, so does $p(\ell_{n+1})$.

(3) So long as $p_{n-1} \ll \mathcal{E}$, $\gamma^{(n)} \simeq 1/\mathcal{E}$, but as p_{n-1} builds up towards unity the enhancement factor $\gamma^{(n)}$ also tends to unity.

To take a concrete case we illustrate in the accompanying figure p_n and $p(\ell_{n+1})$ as functions of n for the particular choice $x = 0.01$, $\mathcal{E} = 0.1$.



3. Quantitative Predictions

We can apply this analysis to the important case of quantitative predictions.[5] Suppose a theory T predicts correctly an experimental result which is known to an accuracy of n significant figures. We assume as part of our background knowledge that the order of magnitude of the result is known, i.e. we disregard the prediction of zeros occuring before or after the significant figures in the experimental result. As a concrete illustration we cite the theoretical predictions by quantum electrodynamics of the anomaly in the hydrogen spectrum known as the Lamb shift and the anomaly in the magnetic moment of the electron. For example the latter is now known experimentally to be $(1159652 4 \pm 2) \times 10^{-10}$ Bohr magnetons[6] whereas the theoretical prediction is $(11596524 \pm 6) \times 10^{-10}$ Bohr magnetons.[7] We thus have remarkable agreement to seven significant figures. Clearly if we regard the prediction of each significant figure as an independent event then the appropriate value to take for ε is 0.1 since a false theory would have ten equal possibilities for filling in each digit. The question of what value to take for x is somewhat arbitrary. In agreement with Schaffner we do not follow a purely logical approach and set x = 0. For our purpose x reflects the scientist's confidence in the new theory T. One could argue that a scientist would not spend great efforts developing the consequences of a theory he did not believe in. By analogy with the situation in the Bayesian analysis of significance testing (see for example Redhead ([1974])) we could take x = ½. Perhaps more realistically we should take x around 0.01 and adopt the sociological rule that unless a scientist has a one per cent level of confidence in the truth of his theory he would not seriously investigate it[8]. With this choice of parameters we see that the build-up of confidence in a theory which makes correct quantitative predictions, as the accuracy of the experiment increases, would be illustrated by the graph for p_n in the figure.[9] Of course $p(e_{n+1})$ is only given by our analysis so long as other factors which could potentially influence the results are known not to be significant. For a certain value